8-1-1958

Limit Distributions of Failure Time for Series — Parallel Systems

Roy Radner

COWLES FOUNDATION DISCUSSION PAPER NO. 55

Limit Distributions of Failure Time for Series - Parallel Systems

R. Radner

August 8, 1958

Limit Distributions of Failure Time for Series - Parallel Systems*

R. Radner

0. Introduction

As systems become more and more complex, the task of computing system reliability from detailed and complete information about the reliability of the components, and about the structure of the system, becomes more and more overwhelming. In such a situation one is led to look for methods of predicting system reliability on the basis of only limited information. Two approaches suggest themselves. First, it would be of interest to know how system reliability depends upon certain average properties of the system, e.g., the average failure time distribution for the class of components used in the system. Secondly, an asymptotic theory might suggest which features of a system become relatively of greatest importance as complexity and size increase. These two approaches are combined in the present paper, to yield a description of the distribution of failure times for a certain type of large series-parallel system, in terms of statistical properties of the structure of the system and of the reliability of the components.

The main result is that as such a system gets more and more "complex" (in a sense that is made precise), the distribution of failure times typically becomes more and more concentrated around a number of discrete time points. The number of these time points depends only upon the structure of the system, but not upon the failure time distributions

---

* Research supported by Boeing Airplane Company, and by the Cowles Commission for Research in Economics under Contract Nonr - 358(01), NR 047-006 with the office of Naval Research. This paper was presented at the New York University - Industry Conference on Reliability, June 11, 1958.

of the components; the location of the limiting time points depends upon both factors. A slightly modified result comes out in the case in which both the components and the connecting links in the system are unreliable; in both cases one is able to calculate explicitly the limiting failure time distribution. Several examples are given.

1. Type of System Considered

The type of system to be considered in this note is characterized by the following properties:

(a) The system is built up of a number of individual components that are connected in a combination of series and parallel arrangements.

(b) Each component can, at any one time, be in only one of two states, "good" or "failed", and once it has failed, it cannot return to the good state. Its time of failure is a random variable that is distributed independently of the failure times of other components. The failure time distributions of different components need not be the same, however, and can be of any form.

(c) The system as a whole, too, can be in only one of two states, good or failed. Provided there is "redundancy" in the system, it is possible for a number of components to fail without causing the system as a whole to fail.

Formally, a <u>system with M components</u> is a function f of M variables $c_1, \ldots, c_M$, such that both the variables and the function can take on only the values 0 or 1. The values 0 and 1 represent the states "failed" and "good", respectively.

One schematic way of representing such a system as a "network" is the following: Each component is assumed to be connected to one or more other components, each connection having a definite direction. This pattern of connections determine a network; it is further assumed that there are no closed loops in this network, i.e. there is no directed path along the connections that starts from a component and eventually returns to it through other components.

A component is called an entrance if there is no connection leading into it from another component; likewise, a component is called an exit if there is no connection leading out from it to another component. When a component fails, all the paths through it are to be thought of as broken at that point. The whole system is good as long as there is an unbroken path from some entrance to some exit; if no such path exists, the system has failed.

Figure 1 shows a system that is still good. Individual components are represented by circles; open circles are good, and filled ones have failed. The connections are represented by arrows, and double arrows into and out of the system indicate entrances and exits.

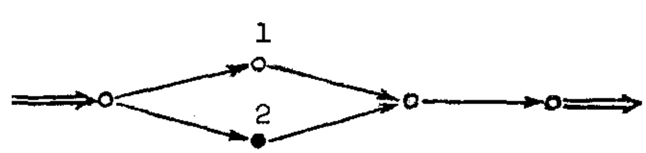


Figure 1.

If, in addition, the component marked "1" in Figure 1 were to fail, then the system would fail, too. However, if component "2" were still good (instead of failed, as marked), then component "1" could fail without causing the system to fail.

A series-parallel system is one that can be built up in stages in the following way. In the first stage, the individual components are grouped into groups of one or more (the size can vary from group to group), and within each group are arranged in series. In the second stage, the series groups are themselves grouped, and are connected in parallel within each group. Call these resulting groups subsystems of order 1. In the next stage, the subsystems of order 1 are grouped, and connected in series within each group, etc. The number of pairs of such stages required to build up the entire system is called the order of the system. Figure 2 shows an example of building up a series-parallel system of order 2.

Formally, a simple series system is represented by the function

$$\prod_{i=1}^{M} c_i \quad ,$$

and a simple parallel system by

$$\sum_{i=1}^{M} c_i \quad ,$$

where the arithmetic used is Boolean, i.e.

$$\begin{cases} 1 + 1 = 1 + 0 = 1 \cdot 1 = 1 \\ 0 + 0 = 0 \cdot 1 = 0 \cdot 0 = 0 \end{cases}$$

A series-parallel system of any order is defined recursively:

(1) $\sum_{j=1}^{M} \prod_{i=1}^{M_j} c_{ij}$ is a series-parallel system of order 1.

(2) If f_{ij} are series-parallel systems of order N, with disjoint sets of arguments (components), then

$\sum_{j=1}^{M} \prod_{i=1}^{M_j} f_{ij}$ is a series-parallel system of order $(N+1)$.

The following points are worth noting:

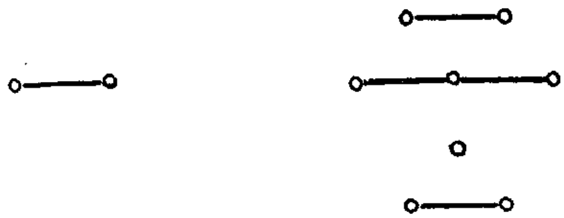
Obviously, one could build up systems by starting with the parallel stage; this might be called a parallel-series system.

Not all networks can be represented as series-parallel or parallel-series systems; Figure 3 shows a simple example.

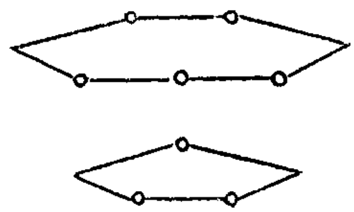
On the other hand, assuming an alternation of series and parallel stages is no less general than allowing any sequence of series and parallel stages, for any two successive stages of the same type could equally well be represented as a single stage.

2. The Statistical Description of the System

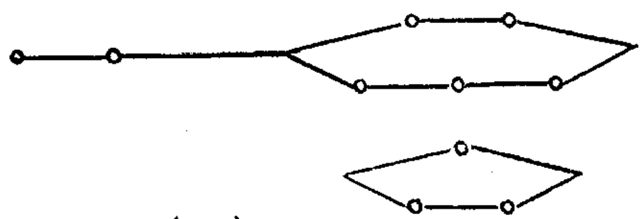
The description of a series-parallel system, as defined in the previous section, can be thought of as having two aspects, (1) the reliability of individual components, and (2) the structure of the network connecting the individual components. This paper is concerned with the implications of certain assumptions about the over-all statistical



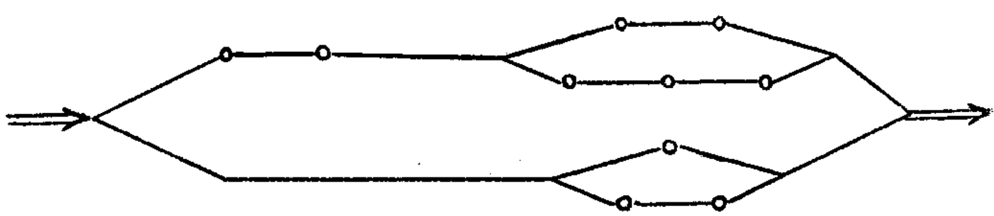
(i)



(ii)



(iii)



(iv)

Figure 2

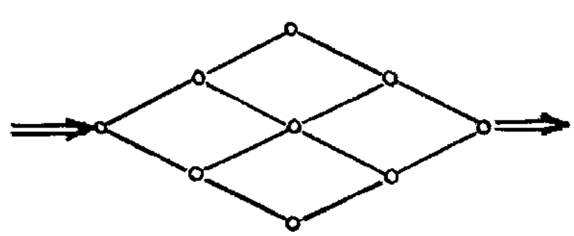


Figure 3

character of these two aspects of the system. The characteristics that are taken as given are:

(1) The average probability of failure by time t, denoted by $A(t)$, for the class of components used in the system; $A(t)$ is defined for every non-negative value of t, and is assumed to be absolutely continuous.

(2) The variance of the probability of failure by time t, denoted by $V(t)$, for the class of components used in the system.

(3) The distribution of the size of groups of components or subsystems connected in parallel.

(4) The distribution of the size of groups of components or subsystems connected in series.

The following assumptions concerning these characteristics are admittedly an over-simplification compared with what one would expect to find in any real system.

It will be assumed that the distribution of series group size is the same for all stages, and likewise for the distribution of parallel group size, but that these two distributions may differ from each other. It is further assumed that the parallel and series group sizes at different stages are independent, and also independent of the failure time distributions of the individual components or subsystems making up the groups.

It will be convenient to characterize the two group size distributions by there respective generating functions, g_S (for groups of components or subsystems in series) and g_P (for groups of components or subsystems in parallel).*

* The generating function of the probability distribution of a random variable N that takes on only non-negative integral values is defined by

$$g(x) = Ex^N = \sum_{n=0}^{\infty} p_n x^n ,$$

where p_n is the probability that N takes the value n .

3. <u>The Distribution of Failure Times for Large Order Series Parallel Systems</u>

Consider now a system of the type just described in Sections 1 and 2. Define the function T on the interval $0 \leqq x \leqq 1$ by

$$(1) \qquad T(x) = g_P[1 - g_S(1-x)] ,$$

where g_P and g_S are the generating functions of the distributions of parallel and series group size, respectively. Let T^N denote the N'th iterate of T (i.e. $T^2(x) = T(T[x])$, etc.); and let X denote the set of values of x between 0 and 1 for which $T(x) = x$. If X is a discrete set, let X_o denote the set of points in X at which $\lim_{N \to \infty} T^N$ is discontinuous; then it can be shown (see Sec. 6) that

(a) as the order of the system increases, the probability distribution of failure time for the system becomes more and more concentrated around a discrete set of time points, which will be called the <u>limiting failure times</u>;

(b) the set of limiting failure times is the set of time points t_o such that $A(t_o)$ is in the set X_o (and may include the time "infinity").

(c) the limiting failure time distribution is

$$A_\infty(t) \equiv \lim_{N \to \infty} T^N[A(t)] \ , \ t \geqq 0 \ ; \tag{2}$$

A_∞ is a step function, with jumps at the limiting failure times;

(d) if the derivative $T'(x)$ is different from 1 at every point of X , then the limiting failure times can be characterized as the solutions t_o of

$$\begin{aligned} A(t_o) &= x_o \ , \\ T(x_o) &= x_o \ , \\ T'(x_o) &> 1 \ . \end{aligned} \tag{3}$$

4. Examples

Example 1. Fixed Group Sizes

Suppose that all series groups have the same size, h, and all parallel groups have the same size, k . Then, regardless of the distributions of failure time of the components used in the system, there will be exactly one limiting failure time for the system, as the order of the system increases without limit.

Figures 4, 5, and 6 show the expected probability of failure as a function of t , for three different pairs (h,k), and for systems of order 1, 2, and 3, under the further supposition that the average probability of failure for the components used in the system is proportional to the time elapsed (with a maximum time of 1000 hours). The figures show that, in these cases at least, the convergence to the limiting distribution is fairly rapid. Thus in the case $h = 4, k = 4$ (Fig. 3) the expected probability of failure occurring between 200 and 390 hours has increased from .19 for the individual components to more than .99 for a system of order 3; in this case the limiting failure time is 278 hours.

Comparison of the three figures shows how the limiting failure time increases when the series group size is reduced, keeping the parallel group size fixed; and how the limiting failure time decreases when the parallel group size (redundancy) is reduced, keeping the series group size fixed.

It is interesting that, if one increases both parallel and series group sizes in the same proportion, the effect is to <u>decrease</u> the limiting failure time (as can easily be shown). This indicates that, <u>if series group sizes are to increase by some factor, then to achieve the same reliability, parallel group sizes must be increased by an even larger factor</u>.

For this example, the transformation T defined by equation (1) is easily found to be

$$(4) \qquad T(x) = [1 - (1 - x)^h]^k$$

<u>Example 2. "Poisson" Group Size Distributions</u>

Suppose that the distribution of series group sizes M_S is such that $(M_S - 1)$ has the Poisson distribution with mean m_S, and that for parallel group sizes, $(M_P - 1)$ has the Poisson distribution with mean m_P. (Note that the average series group size is $(m_S + 1)$ and the average parallel group size is $(m_P + 1)$.) In other words,

$$\text{(4.2)} \qquad \Pr(M=k) = \frac{e^{-m} m^{k-1}}{(k-1)!} \;, \quad k = 1, 2, \ldots$$

with $m=m_S$ for series group size, and $m=m_P$ for parallel group size. Three cases can be distinguished.

Case 1. $(1+m_S)e^{-m_P} < 1$ and $(1+m_P)e^{-m_S} < 1$.

In this case there is exactly one limiting failure time, which is greater than zero. Figure 7 compares the average failure time distribution of a system of order 1 for this case of "Poisson" group size ($m_S + 1 = m_P + 1 = 4$) with that for the corresponding case of fixed group size ($h=k=4$), under the further supposition that the average probability of failure for the components used in the system (the straight line marked "N=0" in the figure) is proportional to the time elapsed, with a maximum time of 1000 hours. The figure shows that the limiting failure times in the two cases are fairly close together, but that convergence to the limiting distribution will be more rapid in the case of fixed group sizes.

Figure 8 shows the average failure time distributions (for the present case) of systems of order 1 and 3, under the supposition that the average probability of failure for the components used in the system is exponential, with a mean failure time of 1 (this last is shown by the curve marked "N=0").

Case 2. $(1 + m_S)e^{-m_P} > 1$.

In this case $t_o = 0$ is the single limiting failure time; in other words, as the order of complexity increases, the entire distribution of failure time becomes more and more concentrated around zero. Figure 9 shows the average failure time distributions for systems of order 1 and 3, when the average failure time distribution for components is exponential with mean 1, and $m_S + 1 = 4$, $m_P + 1 = 1\ 1/2$.

Case 3. $(1 + m_S)e^{-m_S} > 1$.

In this case the only limiting failure time is the maximum possible failure time for the average failure time distribution for components (this is to be interpreted as including the possibility of an infinite limiting failure time if the distribution for components has an infinite range, as in the case of the exponential distribution). Figure 10 shows the average failure time distributions for systems of order 1, 2, and 3, when the average failure time distribution for components is exponential with mean 1, and $m_S + 1 = 1\ 1/2$, $m_P + 1 = 4$. In the limit, for this case, the system becomes "perfectly reliable."

Example 3. "Modified Geometric" Group Size Distributions.

In the previous two examples there was only one limiting failure time, and that appears to be the typical situation within the class of fairly regular group size distributions one might consider. In this example, however, there are two limiting failure times, one equal to zero, and the other equal to the upper limit of the range of the average distribution for the components. Suppose that the series and parallel group size distributions are the same and are given by

$$(5)\qquad \begin{aligned} Pr(M = 1) &= p + q(1 - r) &&, \ q = 1 - p \\ Pr(M = k) &= q(1 - r)r^{k-1} &&, \ k \geq 2 \end{aligned}$$

where both p and r are between 0 and 1 .

Figure 11 shows the average failure time distributions for systems of order 2 and 4, when the average failure time distribution for components is exponential with mean 1, and $p=1/2$, $r = .9$ (for this distribution, the average group size is 5 1/2). The curve for $N=4$ clearly shows a bimodal distribution, with one mode at $t=0$ and the other around $t=4.7$. As the order of complexity increased, the second mode would move out to the right without limit. In the limit, a system would have a probability of approximately .385 of failing as soon as it was put into operation, but if it survived that initial instant, then it would be "perfectly reliable".

Example 4. "Geometric" Group Size Distribution

If, in the previous example, one sets $p=0$, then for both series and parallel group size M , the quantity $(M-1)$ has a geometric distribution. In this case, it can be shown that the average distributions of failure time for systems of all orders are the same as the average failure time distribution for the components! (I have discovered no other case in which this is true.)

General Discussion of the Examples.

It can be shown (see Sec. 6) that the function T defined by equation (3.1) has the property that $T(0)=0$ and $T(1)=1$, hence $x=0$ and $x=1$ are solutions of $T(x)=x$. Therefore, according to 3, it is of interest to know whether the derivative of T is greater than 1 at these two points. These two values of the derivative have the following interpretation (see Sec. 6):

$$T'(0) = (EM_S)\Pr(M_P = 1),$$
$$T'(1) = (EM_P)\Pr(M_S = 1), \tag{6}$$

where EM_S and EM_P denote the average series and parallel group sizes, respectively. The following four cases are of interest. (For this discussion, τ denotes the upper limit of the range of the average failure time distribution for components, possibly infinity.)

<u>Case 1</u>. $T'(0) < 1$, $T'(1) < 1$.

In this case neither 0 nor τ is a limiting failure time. There is at least one limiting failure time between these two values the case in which there is only one is exemplified by (see Example 1 and Example 2, Case 1).

<u>Case 2</u>. $T'(0) > 1$, $T'(1) < 1$.

In this case 0 is a limiting failure time, but τ is not (see Example 2, Case 2).

<u>Case 3</u>. $T'(0) < 1$, $T'(1) > 1$.

In this case τ is a limiting failure time, but 0 is not (see Example 2, Case 3).

<u>Case 4</u>. $T'(0) > 1$, $T'(1) > 1$.

In this case both 0 and τ are limiting failure times (see Example 3).

The reader should bear in mind that in all four cases above there may be other limiting failure times besides 0 and τ , although the examples for Cases 2-4 did not show this.

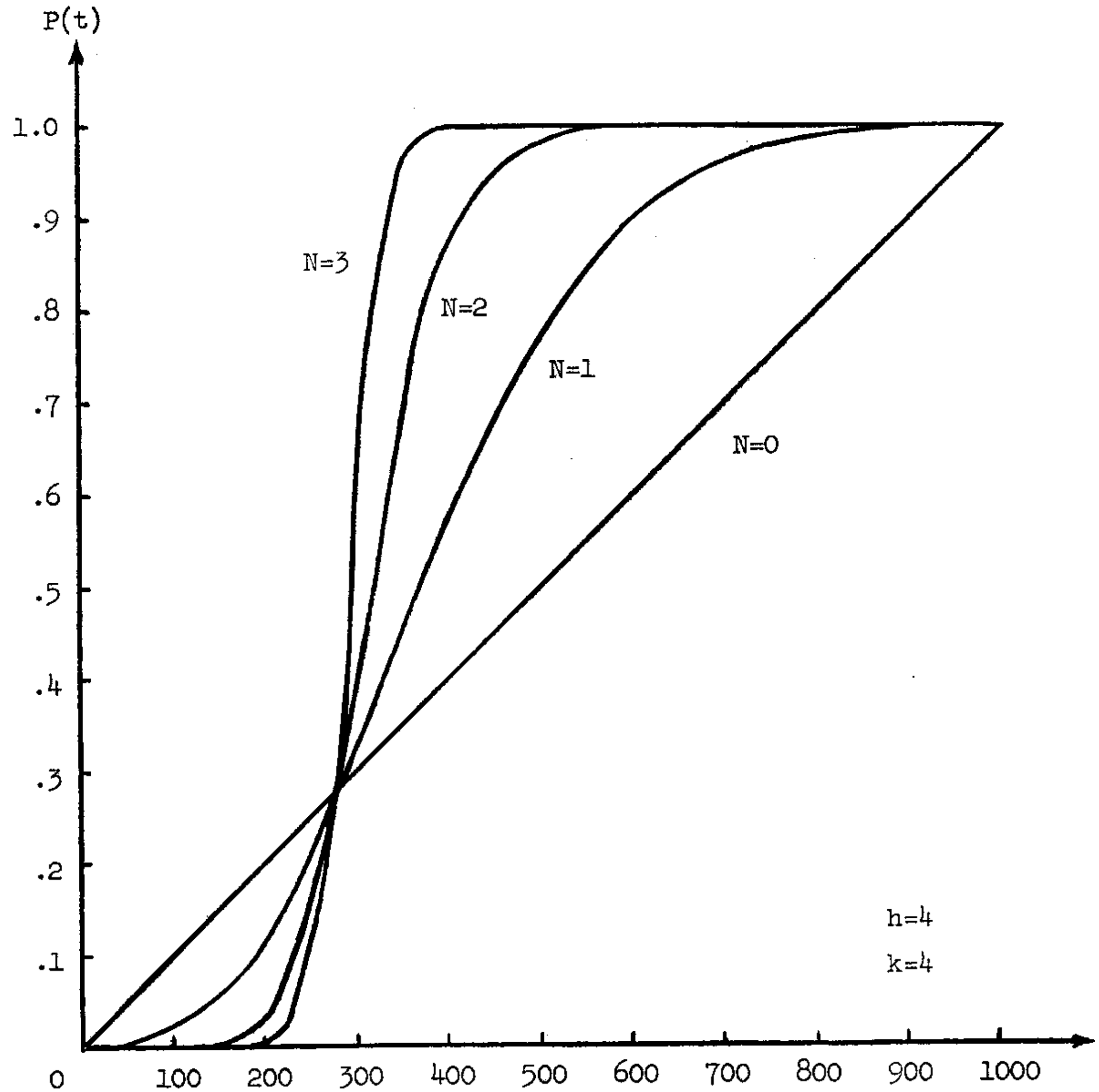

P(t)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
100
200
300
400
500
600
700
800
900
1000
N=3
N=2
N=1
N=0
h=4
k=4

Figure 4

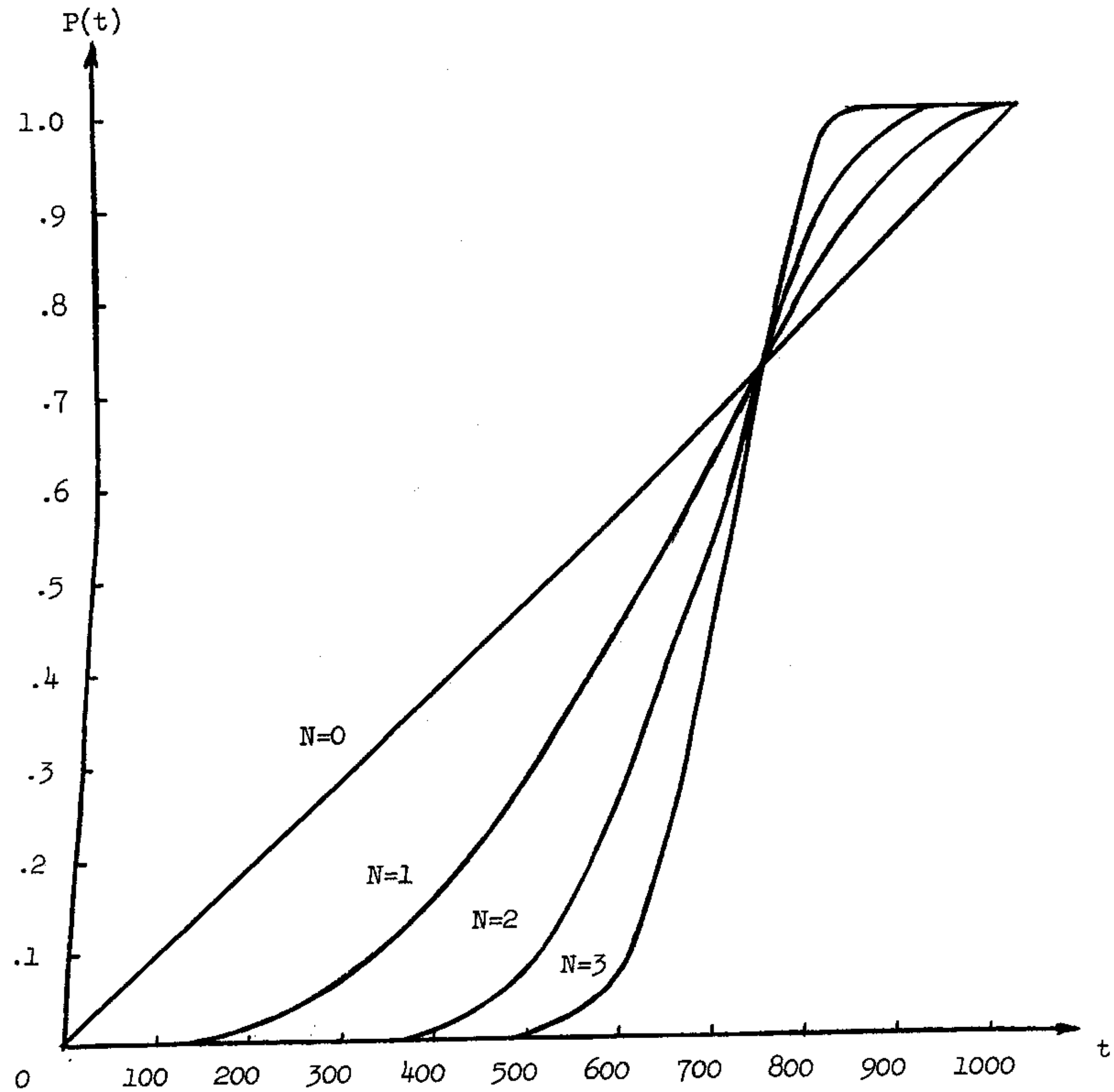

P(t)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
100
200
300
400
500
600
700
800
900
1000
t
N=0
N=1
N=2
N=3

Figure 5

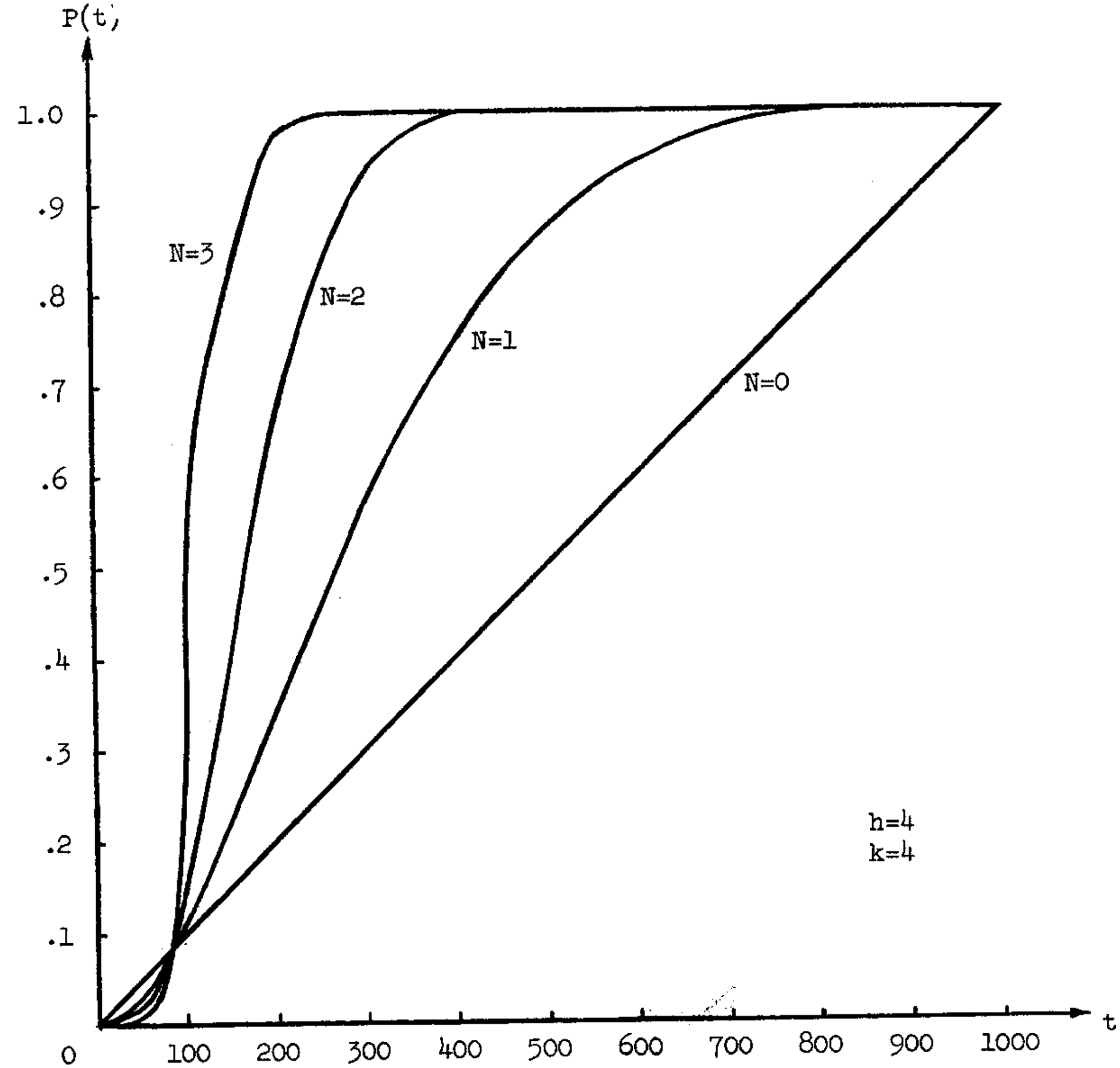
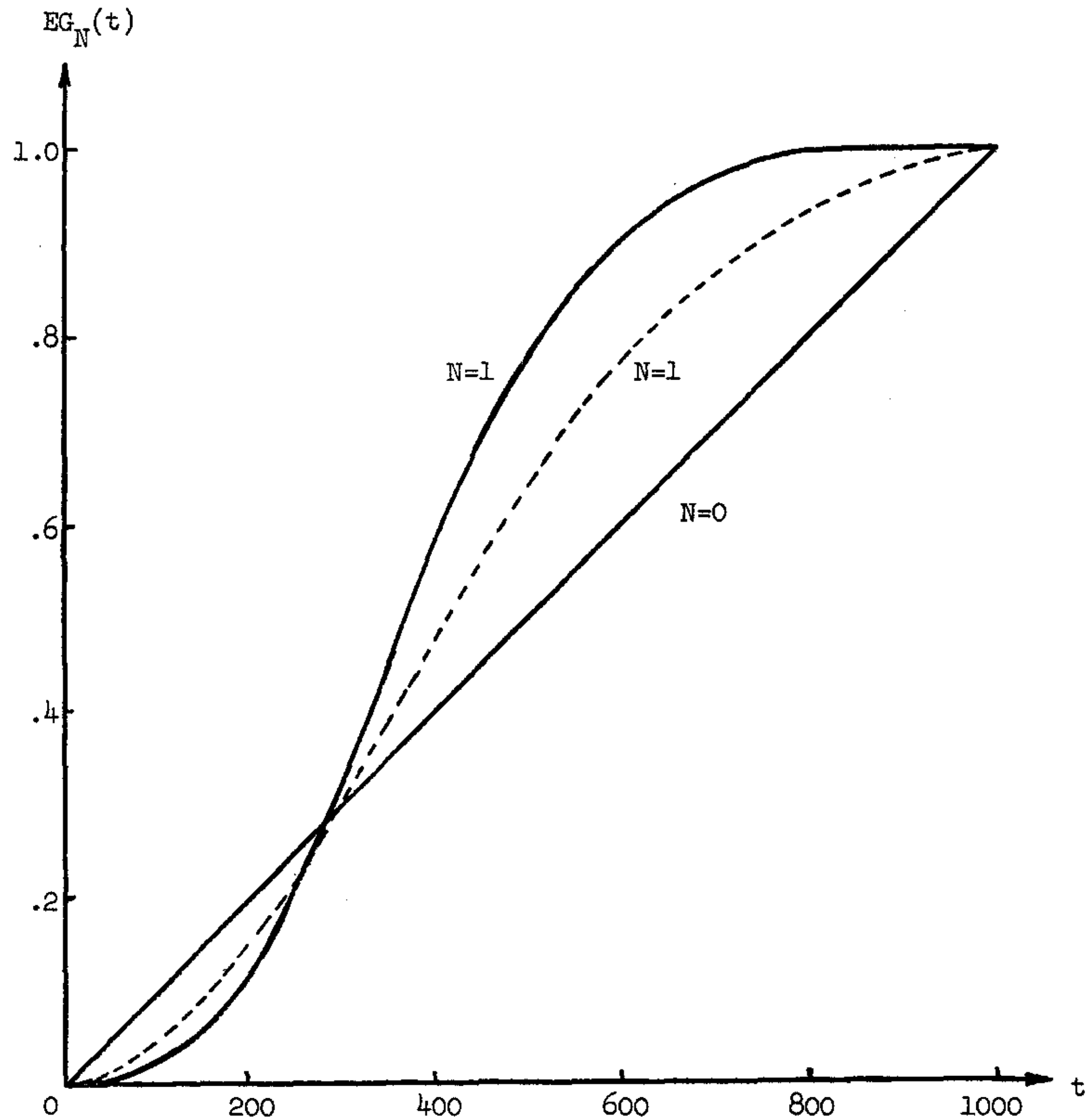

P(t)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
100
200
300
400
500
600
700
800
900
1000
t
N=3
N=2
N=1
N=0
h=4
k=4

Figure 6



- - - - N=1: "Poisson" group size, mean size = 4

———— N=1: Fixed group size, h=k=4

Figure 7

Figure 8. Exponential Components. "Poisson" Group Sizes, Mean Size = 4

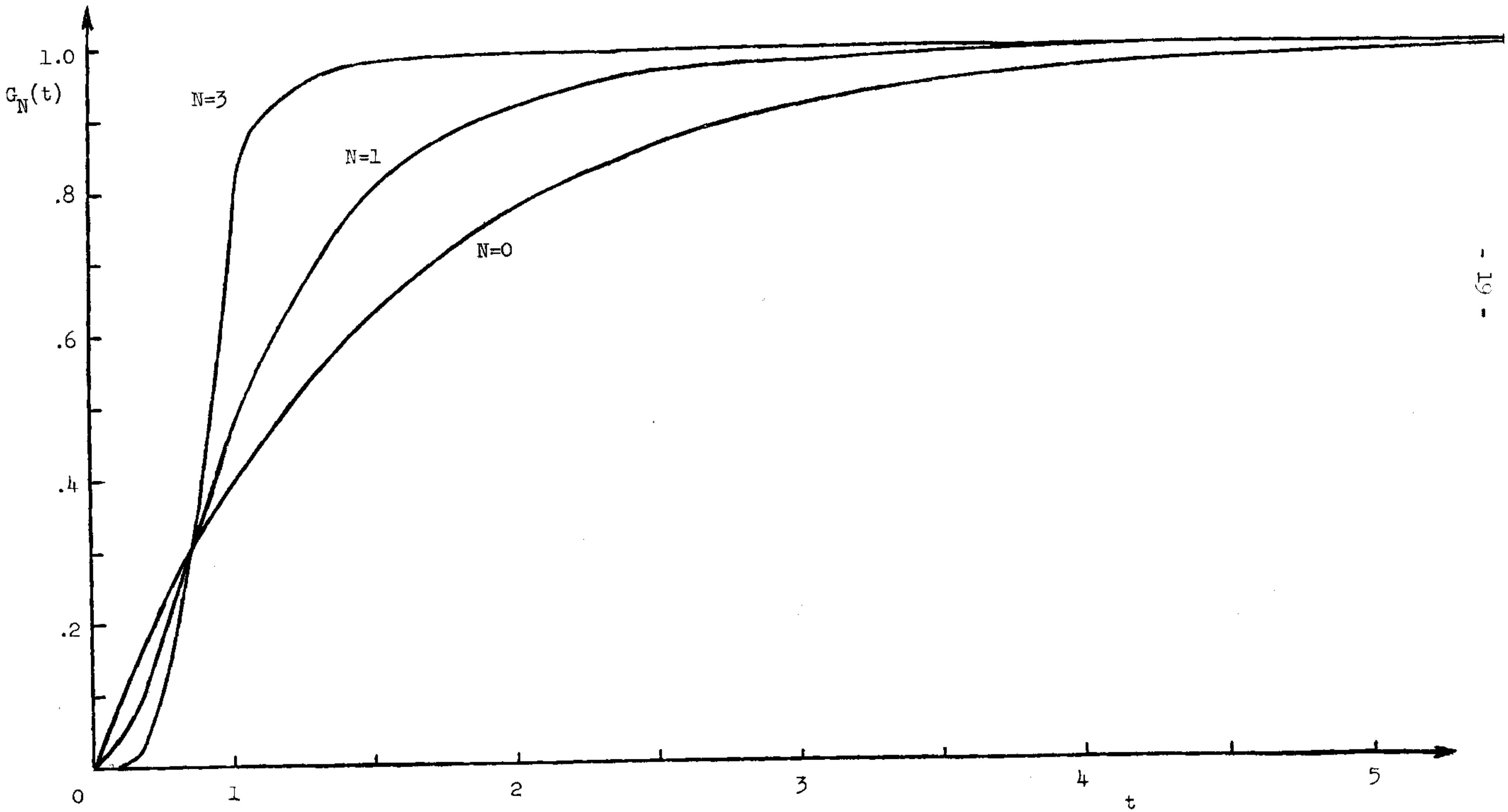


Figure 9. Exponential Components. "Poisson" Group Sizes:
mean for series = 4
mean for parallel = 1 1/2

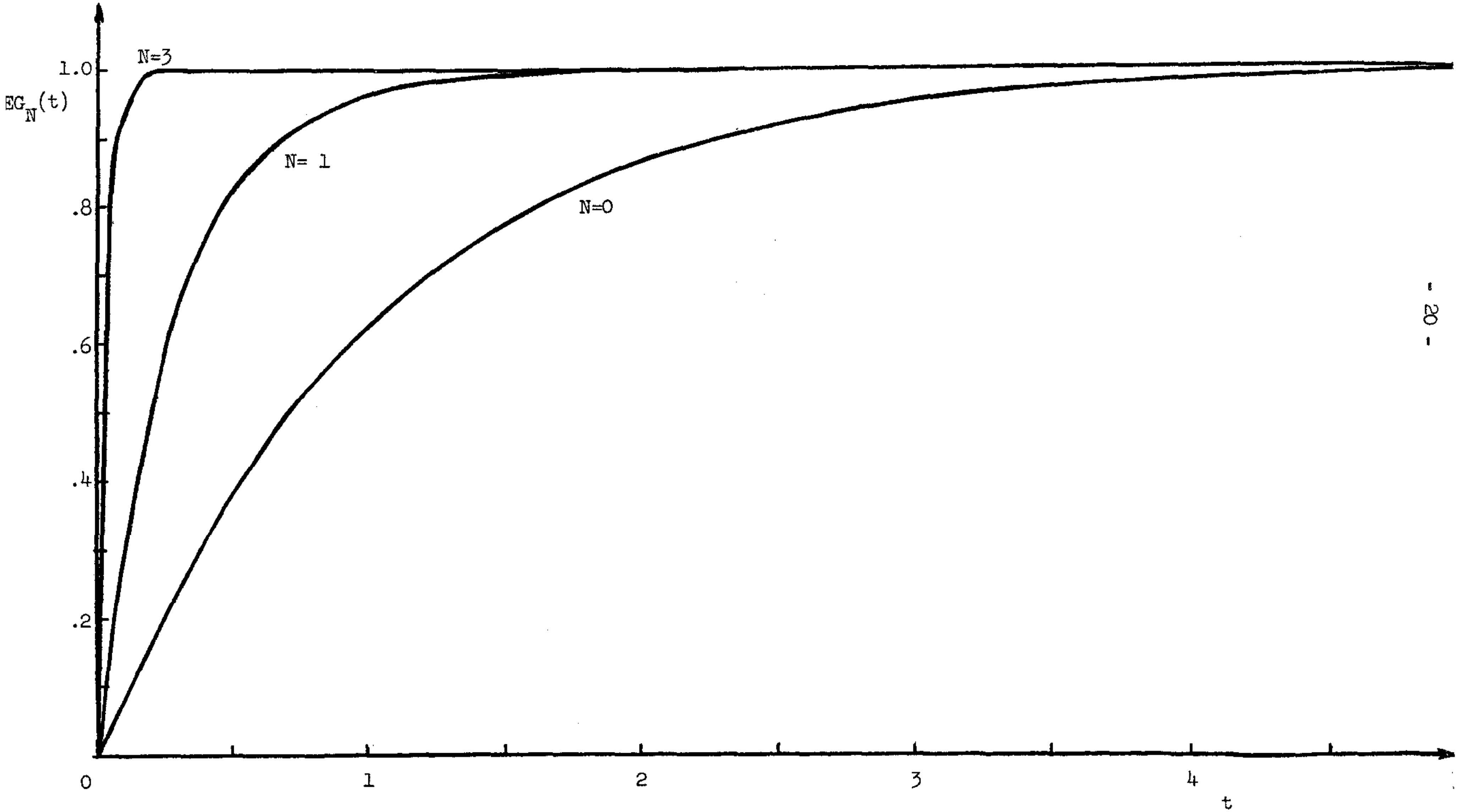


Figure 10. Exponential Components "Poisson" Group Sizes:

mean for series = 1 1/2
mean for parallel = 4

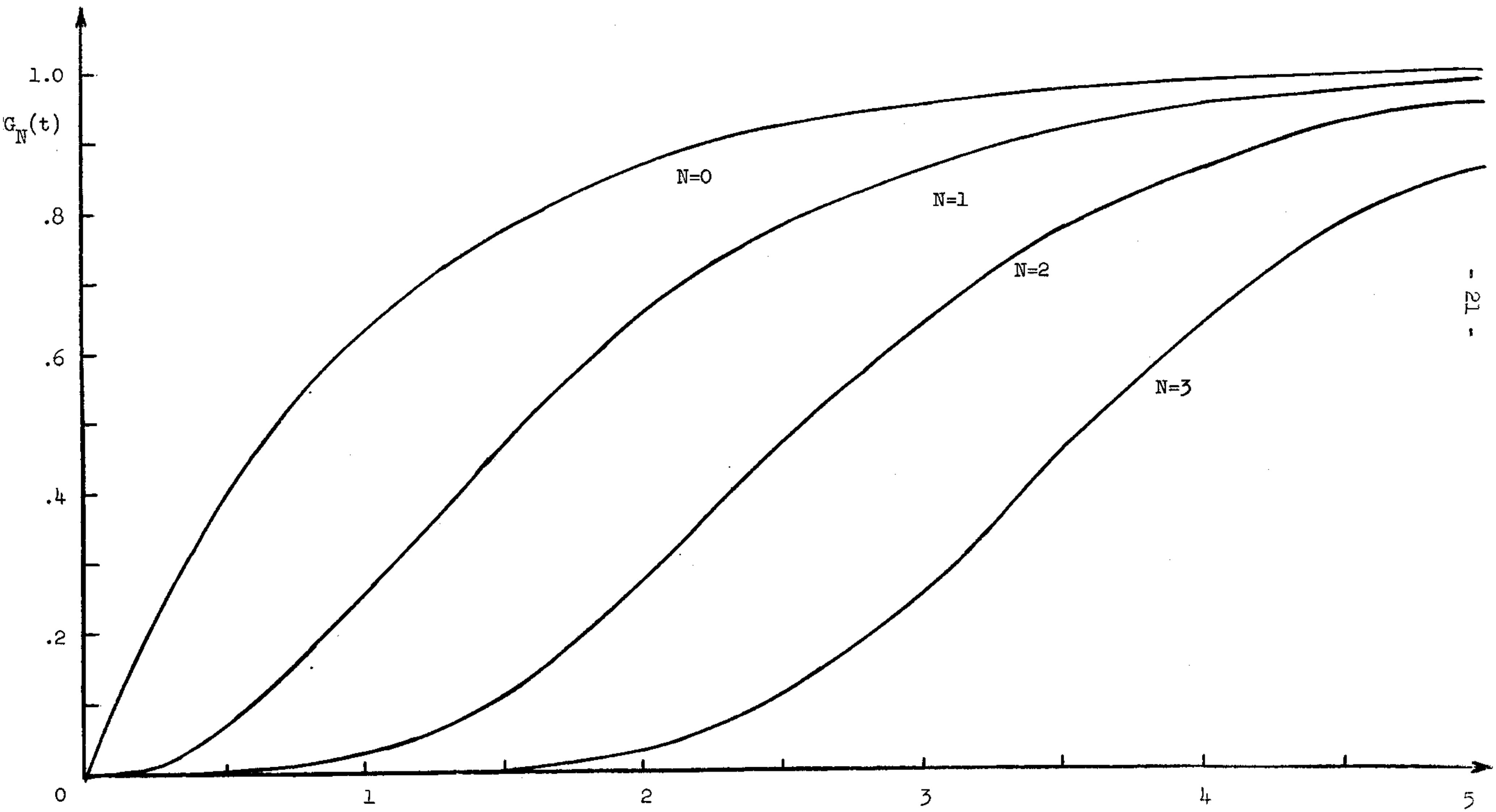
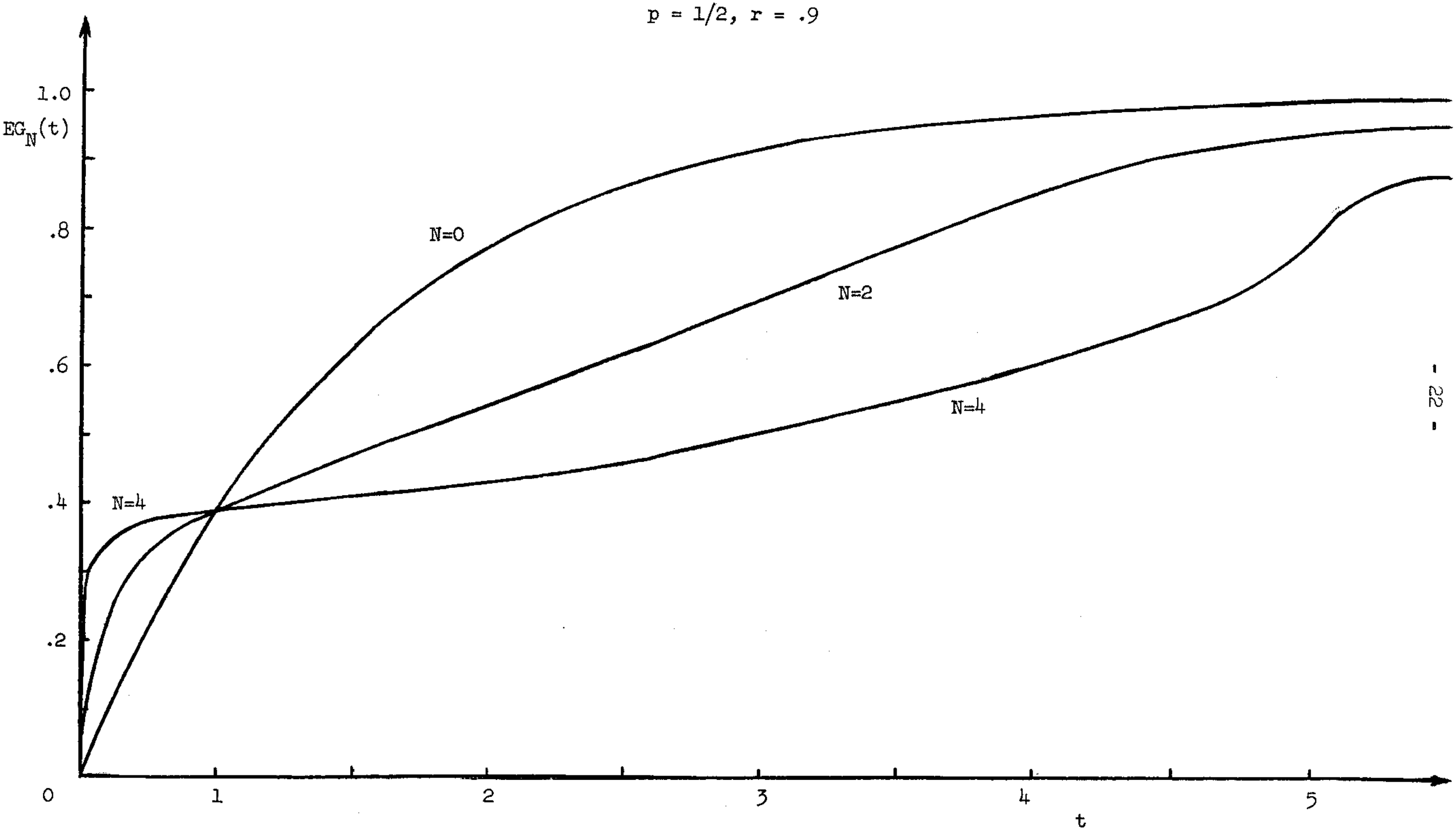


Fig. 11. Exponential Components Modified Geometric Group Sizes:

$p = 1/2,\ r = .9$



5. Extensions of the Results: Unreliable Connections, Mixed Systems.

Thus far in the discussion it has been assumed that the "connections" between the components are perfectly reliable. The simplest concept of a "set of connections" for a single group, series or parallel, is that of an additional component that is placed in series with the group, i.e. the subsystem defined by the group fails if the "set of connections" fails. With this conception of connections it is a straightforward matter to include their effect on the asymptotic reliability of the system. The results are that the limiting expected failure time distribution has jumps, but is not a pure step function, for between the jumps there is a typically small amount of probability (contributed by the unreliability of the connections).

The case of unreliable connections represents a situation in which, at each stage, a certain proportion of subsystems of zero order are mixed in. This suggests the more general situation in which, at each stage, subsystems of various orders are mixed into the groupings at that stage. The variety of possible models here is of course immense, but I can mention one result: if at each stage N, equal proportions of subsystems of all orders less than N are used in the collection from which the new groups are formed, then the asymptotic expected failure time distribution will be the same as in the simple (unmixed) case, but the convergence will be slower.

In practice, one would expect to find that the group size distributions actually change from stage to stage. From the mathematical point of view, such fluctuations could destroy the convergence to a limiting failure time distribution. However, if the group size distributions did not change too rapidly, one might still get a marked approach to a step function at the end of several stages. In any case, the method of analysis used to get the asymptotic results can still be used to get the expected failure time distribution after any finite number of stages.

Finally, I conjecture that the same results (qualitatively) can be obtained for systems built up from basic structures other than simple series and parallel systems. The important mechanism at work here seems to be the repetition of structure (at least in a statistical sense) from one stage to another.

6. Sketch of Proof of Main Theorem.

Let $G_N(t)$ be the cumulative distribution function of failure time for a series-parallel system of order N, generated stochastically in accordance with the assumptions set forth in Section 2. Note that, for every t, $G_N(t)$ is itself a random variable. Recall the definition of the transformation T.

$$T(x) = g_p[1 - g_s(1-x)] \qquad , 0 \leq x \leq 1 .$$

Note that $T(0)=0$, $T(1) = 1$, and T is monotone increasing on $[0,1]$. Let X be the set of solutions x_o of $T(x_o)= x_o$ $(0 \leq x_o \leq 1)$, and let T be the set of solutions t_o of $A(t_o) = x_o$, for some x_o in X. Let T^N denote the N'th iterate of T (i.e. $T^2(x) = T(T[x])$, etc.). The limit $\lim_{N\to\infty} T^N(x)$ exists for all $0 \leq x \leq 1$; furthermore, if the set X of fixed points of T is discrete, then $\lim_N T^N$ will be a step function. In this case, let T_o denote the set of discontinuity points of $\lim_{N \to \infty} T^N[A(t)]$, i.e. the set of t_o such that $A(t_o)$ is a discontinuity point of $\lim_{N \to \infty} T^N$. T_o will be called the set of <u>limiting failure times</u>.

<u>Theorem</u>. Suppose that X is discrete. For any non-negative t_1 and t_2 that are not limiting failure times, and any positive ϵ and δ $(\delta < 1)$, there exists an N_o such that for all $N \geqq N_o$

$$(7) \qquad \Pr\left[\frac{\delta}{2} < G_N(t_1) < 1 - \frac{\delta}{2}\right] \leqq \frac{\epsilon}{2} \ ,$$

$$(8) \qquad \left|\Pr[G_N(t_2) - G_N(t_1)] - [\hat{A}(t_2) - \hat{A}(t_1)]\right| \leqq \delta + \epsilon \ ,$$

where

$$(9) \qquad \hat{A}(t) = \lim_{N \to \infty} T^N[A(t)] \ , \quad t \geqq 0 \ .$$

Inequality (7) says that, as N increases, G_N is more and more likely to be close to a step function with a <u>single</u> jump; (8) says that the probability of that jump being close to any given limiting failure time approaches the magnitude of the jump of $\hat{A}(t)$ at that limiting failure time.

<u>Sketch of Proof</u>.

For fixed series and parallel group sizes, the expected failure time distribution of the simple series-parallel system

$$\sum_{j=1}^{M} \sum_{i=1}^{M_j} c_{ij}$$

is $$EG(t) = \prod_{j=1}^{M_p} \left[1 - \prod_{i=1}^{M_j} (1 - F_{ij}(t))\right] \ ,$$

where F_{ij} is the failure time distribution of component c_{ij}. If M and M_j are distributed according to the assumptions of Section 2, then

$$EG(t) = g_P[1 - g_S(1-A(t))]$$
$$= T[A(t)] .$$

Hence, for a system of order N

$$EG_N(t) = T[EG_{N-1}(t)]$$
$$= T^N[A(t)] .$$

(10) $$\lim_{N\to\infty} EG_N(t) = \lim_{N\to\infty} T^N[A(t)] = \hat{A}(t) .$$

In a similar way one can show that the expected square of $G_N(t)$ satisfies

$$EG_N^2(t) = U[EG_{N=1}(t) , EG_{N-1}^2 (t)] ,$$

where $$U(x,y) = g_P[1-2g_S(1-x)+g_S(1-2x + y)] .$$

From this it can be shown that

(11) $$\lim_{N\to\infty} EG_N^2 (t) = \hat{A}(t) ;$$

hence, by (10) and (11), the limiting variance of G_N is given by

(12) $$\lim_{N\to\infty} \text{Var } G_N = \hat{A}(t)[1-\hat{A}(t)] .$$

Equations (10) and (12) together indicate that for large N, $G_N(t)$ is close to being a binomial variable with probability $\hat{A}(t)$ of taking the value 1, and probability $[1-A(t)]$ of taking the value 0. Equations (7) and (8) constitute a precise statement of this; by using Tchebycheff's theorem they can be shown to follow from (10) and (12).